

## 第七章-逆色调映射

HDR 技术近年来发展迅猛，在未来将会成为图像与视频领域的主流。当前 HDR 内容非常短缺，限制了 HDR 视听节目的广泛应用。逆色调映射 (Inverse Tone Mapping) 应运而生，它是一种用来将 SDR 源信号转换为 HDR 源信号的技术，可以应用于生产端或终端设备，在一定程度上实现对现有 SDR 节目的 HDR “还原”及向上兼容。本系列中，我们将会详细分类介绍逆色调映射算法。分为两个部分：（一）逆色调映射概述及一些经典算法，包括全局算法，分类算法以及拓展映射算法；（二）介绍最近的研究趋势，特别是基于机器学习的逆色调映射算法。

### 7.1 逆色调映射概述

HDR 技术说是继 4K 之后视频领域的又一次重大技术突破，不仅提高图像的动态范围，同时也提供了更宽的色域，并提升了画面在高光和暗部的细节信息，带给观看者更加真实的视觉体验。但作为 HDR 产业链的核心，内容短缺的问题仍然是 HDR 普及和成熟的重要限制因素。众多电视厂商将 HDR 技术作为其产品规划中的重要技术，主要是在硬件方面进行提升。而由于 HDR 内容制作对于设备要求较高，对采集、调色、显示等多方面都需要较高的质量要求，内容制作的复杂度和周期长等因素，导致了 HDR 内容的匮乏。让内容跟上硬件发展的速度，让内容不再限制整个 HDR 产业链的发展依然是当前 HDR 发展完善的重要任务。

如果能充分利用已有的大量 SDR（部分文献中使用 LDR，两者意义一致，本文中均使用 SDR）视频源，将其转换为 HDR 内容，那么不仅可以改善 HDR 片源不足的问题，同时也可以有效利用已有的资源进行内容重置。为此，逆色调映射方法 (inverse Tone Mapping Operator, iTMO 或 reverse Tone Mapping Operator, rTMO) 应运而生。通常，也将该 SDR 转化 HDR 过程的算法称为扩展算子 (Expand Operator)，它可以被定义为一个 SDR 图像  $I$  的函数  $g$ ：

$$g(I) = D_i^{w \times h \times c} \longrightarrow D_o^{w \times h \times c} \quad (1)$$

其中  $D_i \subset D_o$ ， $w$ 、 $h$ 、 $c$  分别是图像的高度、宽度和通道数， $D_i$  是 SDR 图像，在 8 比特量

化的情况下， $D_i = [0, 255]$ ； $D_o$  是 HDR 图像，在使用单精度浮点数时，

$D_o = [0, 3.4 \times 10^{38}] \subset \mathbb{R}$ 。

由于在图像的过曝光和欠曝光部分存在大量的信息缺失，因此从 SDR 的单曝光图片，通过逆色调映射算法进行 HDR 内容的重构是一种病态问题 (ill-posed)，因为这需要算法预测或估计出过曝光或欠曝光部分缺失的信息是什么。因此，过曝与欠曝区域信息的重构，是逆色调映射技术中，需要着重设计的部分。一般情况下，逆色调映射算法将 SDR 内容变换为 HDR 内容需要经过以下几个步骤：

1. 线性化：图片通常已经经过 gamma 校正，也就是说其电平信号与记录的真实场景中的辐射率 (radiance) 并不是线性关系。为了后续处理方便，会先将图像进行线性化处理。

2. 动态范围扩展：该步骤是逆色调映射算法的核心部分，动态范围扩展方式决定着映射后图像的最终质量。对高光部分进行扩展，对低光部分进行压缩，可以达到目的。

3. 过曝/欠曝区域重建：存在过曝或者欠曝现象的 SDR 图像相应区域会损失很多细节，因此需要通过一定的方式恢复或者重建该区域已获得更多的细节。

4. 去噪：逆色调映射算法处理和存储使用的是 16 比特的浮点数，因此需要通过去噪消除量化或压缩而引入的噪声。

5. 颜色校正：进行逆色调映射处理后图像的饱和度会降低，因此需要进行颜色校正让色彩表现 (color appearance) 和 SDR 源相一致。

需要说明，第一步线性化图像非常重要，在线性空间中操作图像的主要原因是对亮度和色域拓展有更大程度上的控制和可预测性。在一个未知的空间中，很难预测扩展如何表现，此外为了恢复场景属性，例如平均值、几何平均值、标准偏差等，需要精确估计场景的辐射度量。

当可以得到拍摄设备信息时，线性化可以通过设备的 CRF 来计算，但当图像储存为 RAW 格式时不需要这个操作，因为 RAW 格式本身就是存储为线性格式。然而，一般情况下我们并不能得到拍摄设备的 CRF，并且视频和图像都是通过 8bit 非线性储存，这种情况下我们需要估计 CRF 来实现线性化操作。一般实现这种线性化的方法，最常用的是 Gamma 校正，除此之外还可以使用辐射校准等。

### 7.1.1 逆色调映射中的颜色转换

SDR 图像与 HDR 图像的区别不仅体现在亮度范围上，也体现在色域差别上。常见的 SDR 图像/视频的颜色空间符合 BT. 709 标准中规定的色域，而 HDR 图像/视频的颜色空间则符合 BT. 2020 标准中规定的色域。这种色域的拓展，使得 HDR 图像/视频在可见光中的颜色覆盖率提高了近 40%。因此，现今的逆色调映射算法不仅仅要对图像的亮度信息与对比度进行增强，还需要进行色域映射。传统的逆色调映射算法很少涉及颜色的处理，忽视这点可能引起增强后的图像/视频出现颜色偏移等现象。

逆色调映射算法中色域转换的常见方法是色彩校正 (Color Correction)，这种方法可以使映射后得到的 HDR 图像获得视觉上与原生 HDR 接近的效果；另一种方法是使用色彩或图像外观模型进行色域转换，一般这种模型都是基于人类视觉系统的特性来实现的，但是这种方法能否准确地预测复杂场景或任意观看条件下的转换参数是值得商榷的。

随着色域的标准化，颜色空间由 BT. 709 向 BT. 2020 的转换过程也被纳入了 ITU 的标准之中，ITU 发布的色域转换建议书 BT. 2087 [25] 中提出了由 BT. 709 到 BT. 2020 的色域转换流程，BT. 2407 [26] 建议书中提出了一种从 BT. 2020 到 BT. 709 色域转换的方法。BT. 2087 中提到的转换方法主要应用于解决 BT. 709/HD 电视内容在 BT. 2020/UHDTV 上播放的兼容性问题，方法是基于光电转换函数和电光转换函数及其反函数，具体转化方法可以参照 ITU 发布的文档。在逆色调映射过程中的颜色转换，可参考 BT. 2087 建议。此外，更先进的色域映射算法考虑如何高效利用图像的空间属性。在空间色域映射中，不再存在从输入色域到输出色域的一对一颜色转换，而是考虑图像的局部背景，使得算法不仅可以准确地再现颜色，而且还能同时再现纹理和局部对比度，其主要方法有两类，一种是将传统映射过程中丢失的信息迭代添加回到图像中；另一类方法，最小化刻画色域映射图像和原始图像之间感知误差的损失函数，一般通过机器学习的方法实现。

逆色调映射算法，可以分为三大类算法：全局算法，分类算法以及拓展映射算法。在本文的后续部分当中，将介绍这三大类算法中的代表性方法，以及最新的基于机器学习的逆色调映射方法。

## 7.2 逆色调映射算法

### 7.2.1 全局模型

全局模型顾名思义，就是对 SDR 内容中的所有像素点都使用相同的扩展方程，该扩展方程对全局有着相同的操作。

#### 1. 幂函数模型

这种方法是最早提出的逆色调映射算法之一，最初由 Landis [1]提出主要从图像重新获取基于幂函数的数字三维模型，主要考虑到幂函数良好的范围拓展能力，用于将小范围的亮度拓展到高动态范围，亮度扩展可以定义为：

$$L_w(x) = \begin{cases} (1-k)L_d(x) + kL_{w,\max}L_d(x) & L_d(x) \geq R \\ L_d(x) & otherwise \end{cases} \quad (2)$$
$$k = \left( \frac{L_d(x) - R}{1 - R} \right)^\alpha$$

其中  $R$  表示扩展的阈值，原文章中使用的是 0.5； $L_{w,\max}$  表示用户需要在输出图片中展示的最大亮度； $\alpha$  是控制色调曲线伸展的衰减指数。

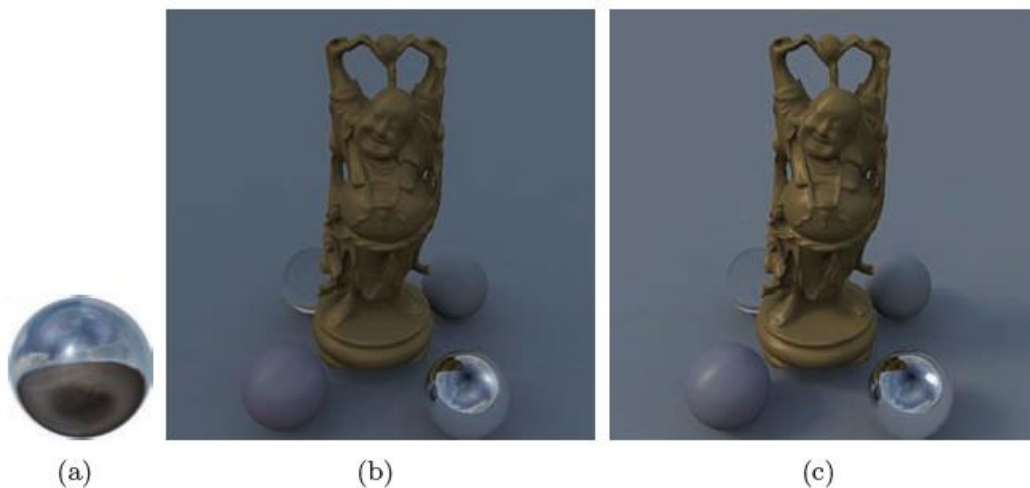


图 7-1 Landis 方法示例：(a)初始 SDR 环境贴图；(b)仅使用(a)的重建图像；(c)使用扩展过的(a)的重建图像。

虽然这种方法为 IBL(image based lighting)生成了合适的 HDR 环境贴图，如图 1 所示，但它可能无法生成可在 HDR 显示器上可视化的高质量图像和视频。这是因为这种方法不处理伪像，例如对压缩或量化产生伪像不作处理，全局的亮度拓展对于整体亮度效果有着较好的改善，但是无法进行细节增强。

## 2. 对 HDR 监视器的线性缩放

为了研究 HDR 显示器支持 SDR 内容的效果, Akyuz [2] 等人进行了一系列有关显示效果的主观测试, 测试实验通过使用杜比 DR-37P HDR 显示器评估色调映射图像、单幅曝光图像和 HDR 图像。实验设计 20-40 岁之间的 22 名测试者, 在所有实验中均使用了从室内到室外、从昏暗到非常明亮的不同光线条件下的 10 张 HDR 图像, HDR 图像亮度约为 5 个数量级, 通过线性缩放处理得到不同动态范围的图片, 以便直接映射到杜比 DR-37P HDR 显示器上。

第一个实验中, 使用各种色调映射算法由 HDR 得到 SDR, 再比较两个图像, 另外也比较了一种自动曝光和测试者在试点研究中选择的曝光, 其中自动曝光是为了最大限度上减少曝光过度 and 曝光不足的像素数量。图像显示在 DR-37P 显示器上, 使用经过校准的 HDR 图像和 SDR 图像匹配 Dell UltraSharp 2007FP 20.1 英寸 LCD 显示器, 测试者的任务主要是将图像按照主观效果排序。对于每个原始测试图像, 测试者必须观看测试图像两秒钟, 这是从不同类型的图像中随机选择的。实验结果表明, 测试者更喜欢 HDR 图像, 作者没有发现色调映射和单次曝光图像在测试者之间的偏好方面存在很大的差异。

第二个实验中, 作者比较了从单次曝光图片扩展得到的 HDR 图像和单次曝光的图像, 为了扩大单次曝光图像, 扩展方法可以表示为:

$$L_w(x) = k \left( \frac{L_d(x) - L_{d,\min}}{L_{d,\max} - L_{d,\min}} \right)^\gamma \quad (3)$$

其中  $k$  表示 HDR 显示器的最大亮度强度,  $\gamma$  表示非线性缩放因子。对于这个实验, 分别生成了  $\gamma$  为 1、2.2 和 0.45 的图像, 此外设置和排名方式与第一个实验相同。结果表明, 选择较亮的曝光扩展图像比 HDR 图像更受欢迎, 当它们具有相同的平均亮度时反之亦然。Akyuz 等人表示平均亮度对比度更受欢迎。最后, 另一个重要结论是  $\gamma = 1$  时产生的线性缩放图像最受欢迎, 这表明线性缩放足以满足 HDR 图像和视频的体验。

作者仅使用高分辨率 HDR 图像, 并且没有产生压缩伪像。虽然这在理想条件下工作良好, 但在实际情况下, 例如使用压缩的电视节目或 DVD 中, 情况可能并非总是如此。在这些情况下, 需要进行更准确的扩展以避免压缩伪像和轮廓的放大。

## 3. 过曝 SDR 图像的 Gamma 扩展

Masia [7] 等人进行了两项心理物理学研究, 以分析扩展函数在各种曝光度水平下的表现。作者随后使用这些实验的结果来开发过曝内容的扩展技术。在这些实验中, 比较了三种逆色调映射算法, 包括 Rempel[8]、线性缩放[2]和 Banterle[13]提出的方法(这些算法后文中将介绍), 并且结果显示随着内容曝光度的增加, 结果下降的趋势。作者还观察到, 逆色调映射函数制造出一些在图像 SDR 格式中可见的伪像, 原因在于算法产生不正确的光照强度值, 并造成空间内容失真。了解逆色调映射函数如何影响最终输出图像的观看效果可能有助于开发更好的算法。

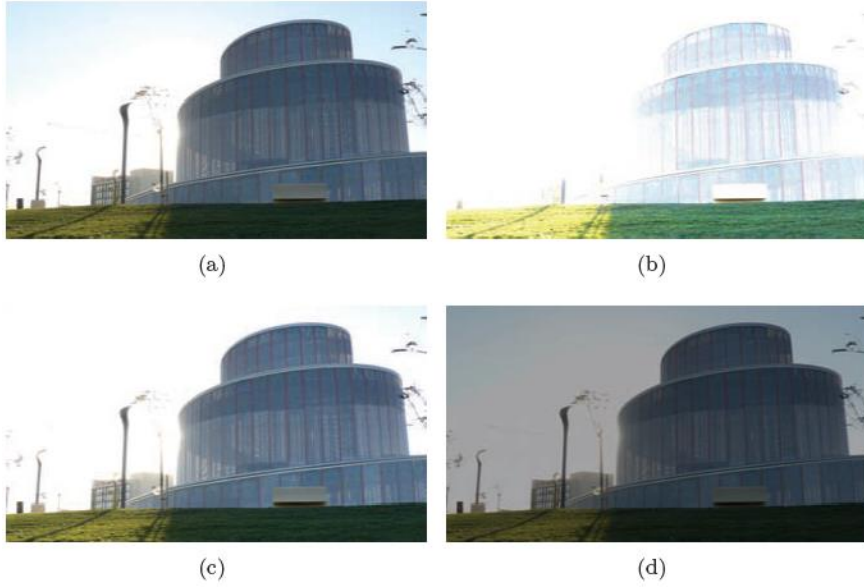


图 7-2 Masia 方法中过曝图片案例：(a) 原始 SDR 图像；(b) (c) (d) 扩展后不同光圈下的图片。

如图 2 所示，作者指出：当图片的曝光度提高时，细节会丢失，这是由于像素点的饱和度增加使得对应的颜色褪色至白色。基于这种发现，他们提出通过将剩余细节突出更加明显来映射 SDR 图像，这不同于常见的逆色调映射算法只是增强饱和的区域。

一种实现这种方法的直观的方式，就是使用 Gamma 扩展，Masia 提出一种自适应的方式，通过输入 SDR 图像的动态内容来确定  $\gamma$  的值，与 Akyuz 的方法类似， $k$  值可以通过以下方式计算：

$$k = \frac{\log L_{d,H} - \log L_{d,\min}}{\log L_{d,\max} - \log L_{d,\min}} \quad (4)$$

其中  $L_{d,H}$ ,  $L_{d,\min}$ ,  $L_{d,\max}$  分别是对数平均数、最小亮度值和最大亮度值， $k$  值是一个统计数据，有助于澄清输入图像是主观黑暗还是明亮。为了自动预测 Gamma 值，开展了一项试点研究，要求用户手动调整一组图像中的  $\gamma$  值，数据是经验拟合关系的线性回归：

$$\gamma(k) = ak + b \quad (5)$$

由拟合得到， $a = 10.44$ ,  $b = -6.282$ 。这种逆色调映射算法的主要缺点之一是它可能无法充分利用动态范围，而且  $a$  和  $b$  的值只能在该组测试图像上正常工作，在一些不属于原始图像集的图像中， $\gamma$  可能为负值。

## 7.2.2 分类模型

### 1. HDR 显示器的高光重现

Meylan [9, 10]提出一种专门用于处理包含镜面高光的图像的逆色调映射算法，是一种典型的分类模型算法。其主要思想是检测图像中的漫反射和镜面部分，并使用不同的线性方程对其进行扩展。

算法的第一步是计算一个阈值  $\omega$ ，用于在亮度通道区分高光部分和漫反射部分，如图 3 所示。首先，图像通过使用一个大小为  $m = \max(\text{width}, \text{height}) / 50$  的滤波器，来计算  $t_1$  作为滤波后亮度通道的最大值，重复使用上述操作可以使用一个  $2m+1$  大小的滤波器来计算  $t_2$ 。其次， $t_1$  被用于在原始亮度上的阈值，用于生成一个 mask，随后使用  $t_2$  对 mask 应用腐蚀和膨胀滤波器。这些可以应用于一些迭代以获得稳定的 mask，一般情况下几次迭代就可以达到效果。此时，mask 中像素值为 1 的点被视为镜面像素，而黑色像素被视为漫反射像素。因此， $\omega$  被作为  $L_d$  中的镜面像素的最小亮度值。

在计算  $\omega$  之后，亮度通道可以通过以下方式扩展：

$$L_w(x) = f(L_d(x)) = \begin{cases} s_1 L_d(x) & L_d(x) \leq \omega \\ s_1 \omega + s_2 (L_d(x) - \omega) & \text{otherwise} \end{cases} \quad (6)$$

$$s_1 = \frac{\rho}{\omega}, s_2 = \frac{1 - \rho}{L_{d,\max} - \omega}$$

其中  $L_{d,\max} = 1$  是由于图像被归一化处理过， $\rho$  表示 HDR 显示器分配给漫反射部分的百分比，这是由用户定义的。

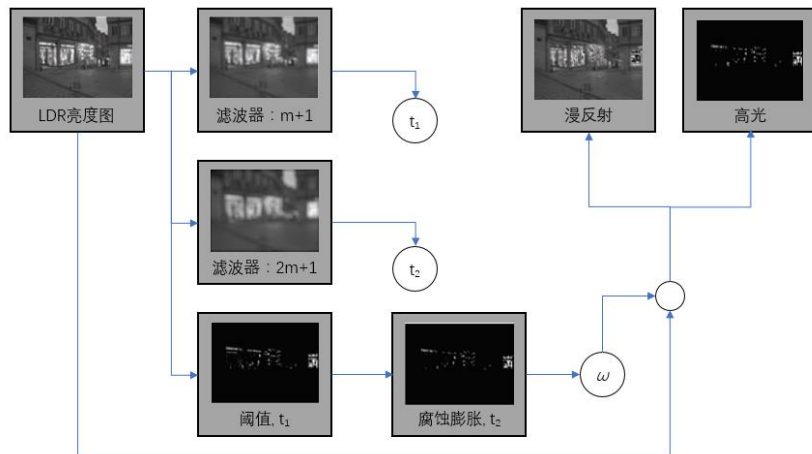


图 7-3 计算最大漫反射亮度值  $\omega$  的计算框图

$f$  的全局应用可能会导致增强高光部分的量化伪像，通过在镜面反射区域应用选择性

率可以减少这些伪像，图 4 显示了算法整体的流程。首先使用  $5 \times 5$  平均滤波器对扩展后的亮度值  $f(L_d(x))$  进行滤波，得到  $f'$ 。随后， $f(L_d(x))$  和  $f'$  使用线性内插和 mask 进行混合，该 mask 是通过使用  $\omega$  对 SDR 亮度进行阈值化处理并应用扩张和  $5 \times 5$  平均滤波器来计算。

最后，作者使用杜比 DR-37P HDR 显示器进行了一系列心理物理学实验，以确定  $\rho$  值。结果显示，对于室外场景，用户偏好较高的  $\rho$  值，这意味着分配给高光的动态范围的百分比很小，而室内场景的百分比则相反。对于具有相同漫反射亮度的室内和室外场景，用户选择较低的  $\rho$  值，因此他们更喜欢将更多范围分配给高光，数据分析表明  $\rho=0.66$  是一个很好的一般估计值。

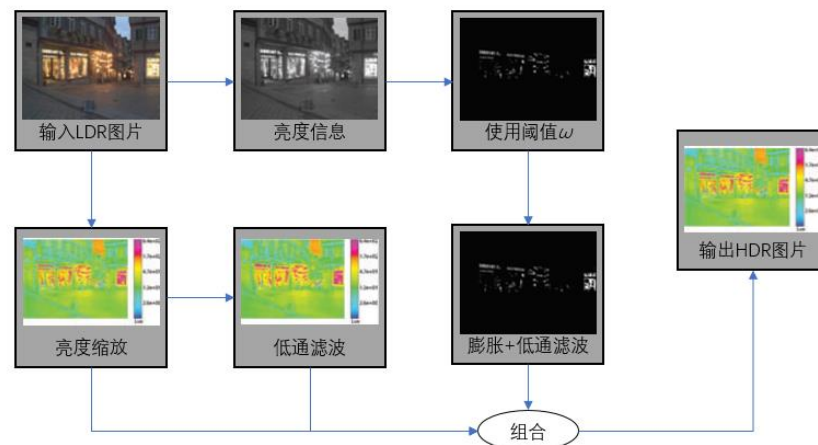


图 7-4 Meylan 方法整体框图

镜面高光相比漫反射高光部分具有更高的亮度，通常在图像采集或者处理阶段往往会进行压缩或者截断，损失了一定的图像细节和对比度。Meylan 通过心理视觉实验发现，对于传统的 SDR 内容，如果直接采用线性 iTMO 进行动态范围扩展，必然会引起图像镜面高光部分泛白，图像饱和度不足，图像整体视觉表现不自然。该算法转为特定服务而设计，在 HDR 监视器上重现高光点，其他任务如增强食品的使用需要更多处理和分类操作，作者的评估实验强调了这一点。

## 2. HDR 显示器的明亮视频特征增强

Didyk [11]等人提出了一个用于增强 SDR 视频亮度的交互式系统，主要用于 DVD 内容的显示和增强。这个系统的主要思想是将场景分为三个部分：漫反射、反射和光源，然后仅对反射和光源进行增强。作者提到，漫反射部分在不产生视觉伪像的情况下难以增强，并且电影制作者可能会向他们展示饱和度，而不是光源和截断反射。该系统适用于非线性值，因为目标是增强而非物理准确性。



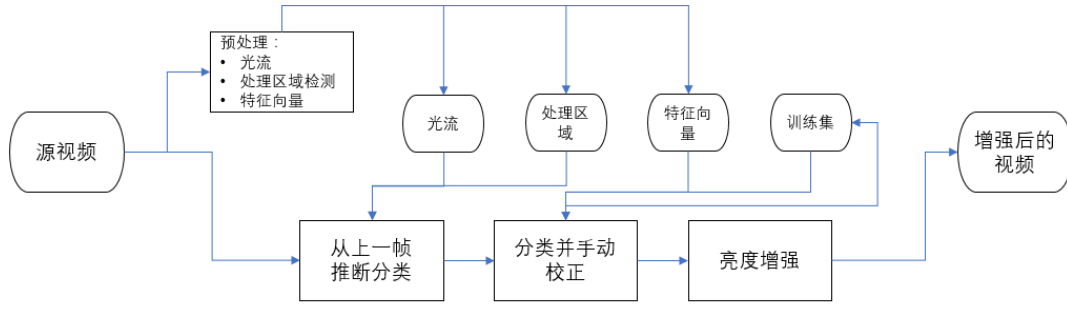


图 7-5 Didyk 方法流程图

该系统由三个主要部分组成：预处理、分类和增强裁剪区域，流程框图如图 5 所示。预处理步骤会生成分类过程中所需的数据，特别是它使用 flood-fill 算法来确定剪切区域，至少有一个通道必须饱和，而且亮度值须大于 222。此外，在此阶段，还会计算光流以及其他功能，如图像统计、几何特征和邻域特征。

分类决定了框架中的亮区、反射和漫反射区域，并依赖于 2000 个人工分类区域的训练集。首先，具有内核  $k(z, z') = \exp(-\gamma \|z - z'\|^2)$  的支持向量机执行区域的初始分类；随后，使用基于欧几里得度量的最近邻分类器，运动跟踪改进了初始估计：

$$d^2((z, x, t), (z', x', t')) = 50 \|z - z'\|^2 + \|x - x'\|^2 + 5(t - t')^2 \quad (7)$$

其中  $z$  表示区域特征， $x$  表示图像中的坐标， $t$  表示帧数。这允许在测试中使用的所有区域达到 12.6% 的分类误差，使用运动补偿跟踪修剪区域进一步减少了需要手动修正的对象的百分比，达到 3%。

通过应用非线性自适应色调曲线来增强修剪区域，这是基于存储在直方图  $H$  中的修剪区域内的偏导数计算的，色调曲线被定义为  $H$  的倒数的直方图均衡：

$$f(b) = k \sum_{j=2}^b (1 - H[j]) + t_2 \quad (8)$$

这里  $t_2$  是修剪区域的最低亮度值， $k$  是限制到最大提升值  $m$  缩放因子（对于亮光为 150%，对于反射为 125%）：

$$k = \frac{m - t_2}{\sum_{j=1}^N (1 - H[j])} \quad (9)$$

值  $N$  是  $H$  中的 bin 的数量，为了避免在增强期间产生轮廓，使用双边滤波对亮度通道进行滤波，将其分离为精细细节和基础层，其在亮度扩展之后合并，该方法是半自动的，因为需要用户的介入。

### 7.2.3 扩展映射模型

扩展映射模型（Expand Map Models）最早是 2006 年 Banterle 在其一篇文章“inverse tone mapping” [13] 中提出的术语，表示在将 SDR 内容扩展为 HDR 的过程中使用一个引导函数指导该过程的转换。该模型在 Banterle 及 Rempel 的文章中均有使用。



## 1. Banterle——使用扩展映射模型进行非线性扩展

Banterle [13]提出了扩展 SDR 内容到 HDR 监视器的基本流程“inverse tone mapping/a framework for inverse tone mapping”，如下图 6 所示。该方法的核心即如何使用反向的 TMO 算子，并结合一定的平滑处理算法重建出过曝区域中丢失的图像信息。

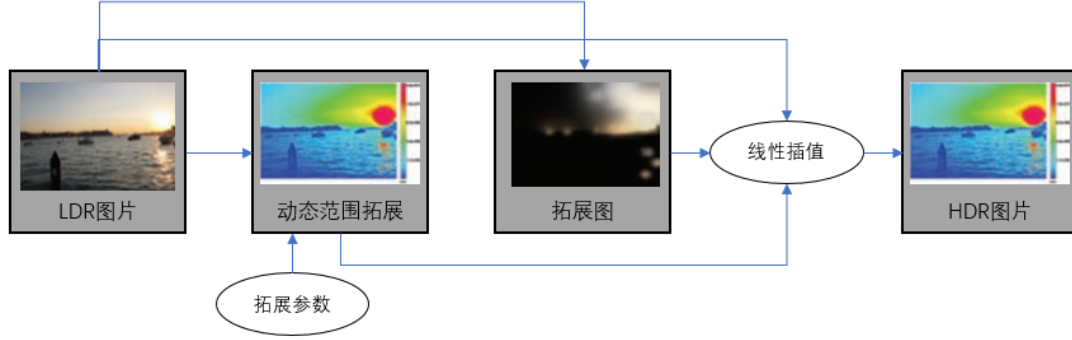


图 7-6 Banterle 方法的流程框图

本方法的逆色调映射方程可以被描述为：

$$L_w(x) = \frac{1}{2} L_{w,\max} L_{white} \left( L_d(x) - 1 + \sqrt{(1 - L_d(x))^2 + \frac{4}{L_{white}^2} L_d(x)} \right) \quad (10)$$

其中  $L_{w,\max}$  表示输出亮度的最大值，以  $\text{cd/m}^2$  为单位， $L_{white} \in (1, +\infty)$  是一个决定扩展曲线形状的参数，与对比度成比例。作者建议让这两个值取等，以此来在限制扩展导致的伪像，同时增加对比度。

在范围拓展之后，计算扩展映射是表示高亮度区域中图像的低频版本的平滑区域，它有两个主要目标，首先是重建图像曝光过度区域丢失的亮度分布；第二个目标是减弱扩展期间可以增强的量化或压缩伪像。扩展映射通过对使用重要性采样生成的样本应用密度估计来实现。最后，扩展 SDR 图像和原始 SDR 图像通过线性插值进行组合，其中扩展映射充当权重加权重。需要注意的是，低亮度值保持原始值，这样可以避免在  $L_{white}$  设置为较高值时对低值进行压缩，否则可能会出现轮廓等伪像。

为了方便后期处理，首先需要把图像进行线性化，然后使用反向的 TMO 算子进行范围扩展。得到线性化的输入图像之后，根据输入图像，利用中位数采样强度估计的方法计算扩展映射图。其包含了图像的高亮区域的低频分量，提取高亮区域的低频分量是用于重构高亮区域中丢失的图像信息。另外，在范围扩展的过程中，量化和压缩过程必然会带来失真。使用扩展映射的另一个目的是为了尽量消除该影响。最后，将扩展映射作为插值权重，将原始的 SDR 图像和范围扩展得到的图像进行线性插值，得到最终的扩展图像。

Banterle 将该模型进行了扩展，使其可以用于自动处理图像和视频，扩展参数需要通过强度估计和试探法确定以重构高亮区域丢失的图像信息。该方法的优势在于使用扩展映射很好的保持了原始图像中的边缘信息，从而减小了最终生成的图像的边缘失真。该方法也有一定的局限性：当图像中的过曝区域过大（大于 30% 左右时），没有足够的信息用于图像区域的重建，不能通过强度估计等方法进行图像信息恢复。

## 2. Rempel——SDR2HDR

由于 Banterle 提出的扩展映射模型在速度方面具有一定的限制，如果不使用高端的硬件支持，很难满足实时性要求。2007 年 Rempel [8]提出了一种基于扩展映射模型的简单方法，该方法可以实时地将 SDR 图像进行扩展得到 HDR 图像。该方法的流程图如下图所示：

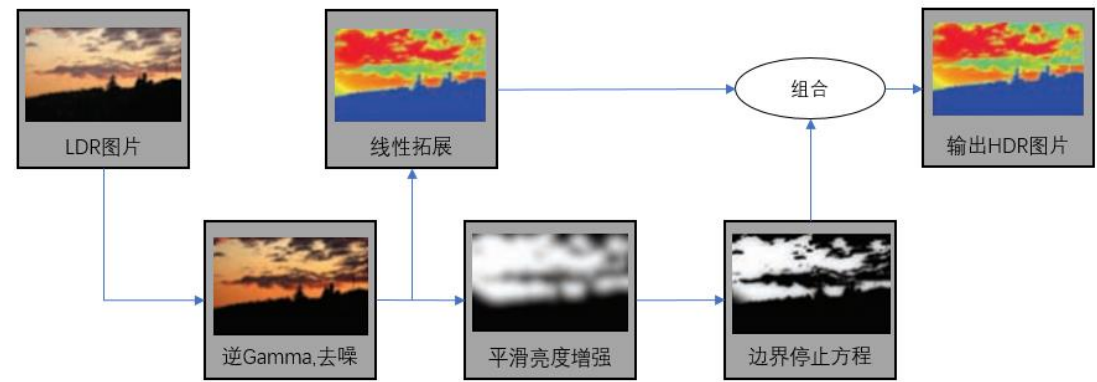


图 7-7 SDR2HDR 方法流程框图

该算法首先对输入 SDR 图像进行去线性化和对比缩放的操作，一般我们看到的 SDR 图片都是经过 Gamma 校正后的图像，在逆色调映射中通常需要将图像还原至线性空间。但是经过实验发现，数码相机拍摄的照片和视频通常还经过了传感器本身的曲线的处理，比常见的 Gamma 校正更为复杂，算法在本步骤的处理就是为了模拟数码相机的反应曲线，使得得到的图片更接近模拟胶片的效果。实现是通过实验验证设定了一个简单的 Gamma 曲线，从而得到了一种视觉上令人满意的效果。

得到线性空间的图像之后，为了适配 HDR 显示设备，将图像的亮度信息缩放到整个 HDR 显示设备支持的亮度范围内。但是这种缩放不能改变过多，否则会造成图像看起来不够自然，最重要的是将高亮部分变得更亮，使其产生类似“发光”的效果。大量实验结果得出，算法设定的对比度大概在 5000:1，亮度范围在  $0.3 \sim 1200 \text{cd/m}^2$  之间。随后，为了去除这种缩放所产生的噪声和量化误差，使用一个简单的滤波操作来处理图像。

算法的第二步是对图像的亮区进行增强，首先将亮区分割出来，算法认为视频中像素值高于 230、图像中像素值高于 254 的区域为亮区，分割出来存为 mask，再获得图像的轮廓图用于之后的处理。

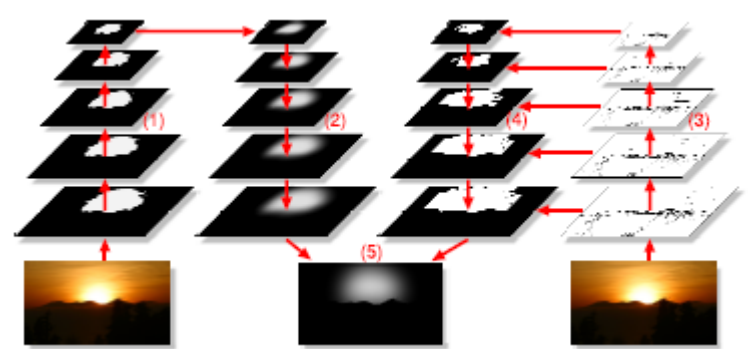


图 7-8 使用图像金字塔进行亮度增强

如图 8 所示，亮度增强是通过图像金字塔进行的。首先是过程 (1) 对图像的亮区 mask 进行下采样，得到光源的信息，然后如过程 (2) 对结果使用最近邻插值上采样，得到平滑后的

亮区光源信息；同时(3)对轮廓图也同样进行下采样处理，得到图像的强轮廓信息，在过程(4)上采样亮度 mask 信息时，使用同分辨率下的轮廓信息，在上采样的过程中，遇到强轮廓就停止采样，从而使得亮区不会产生过量扩散的错误。最后(5)就是将平滑过的亮度信息与 mask 融合，得到最终 HDR 图像的高亮区域亮度信息。

该算法本质上可以算作是一种基于 Expand Map 的方法，主要对 SDR 图像上无法显示的高亮区域进行增强，使得图像在 HDR 显示器上可以更为自然的显示出来。由于算法的计算复杂度较小，可以达到实时处理的效果，这相比之前 Banterle 的实时方法对硬件的要求更低。但是本算法只对高亮区域处理，而对正常亮度区域的增强不明显，在正常场景拍摄的图片中增强效果不明显，且算法没有考虑色域增强，只是简单地增加饱和度来实现更鲜艳的颜色。

## 7.2.4 近年提出的逆色调映射算法

### 1. 适用于各种曝光的高质量逆色调映射

Kovaleski 和 Oliveira [22]在 2014 年提出了一种增强算法，这种方法基于交叉双边滤波，并能够很好的处理曝光不足或过量的情况。

双边滤波是一种非线性滤波器，同时对空间域和亮度范围进行滤波。使用一个高斯方程  $G$  作为下降函数，则双边滤波后的  $I^b$  可以被定义为：

$$\begin{aligned} I_p^b &= \frac{1}{W_p^b} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|I_p - I_q|) I_q \\ W_p^b &= \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|I_p - I_q|) \end{aligned} \quad (11)$$

其中  $I_p, I_q$  代表在图像在位置  $p$  和  $q$  上对应的像素值， $\sigma_s, \sigma_r$  表示空间和距离上的高斯函数标准差， $W_p^b$  是一个归一化因子。将其中的  $G_{\sigma_r}(|I_p - I_q|)$  部分替换为另一幅图片，那么上式就可以得到交叉双边滤波：

$$\begin{aligned} I_p^b &= \frac{1}{W_p^b} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|E_p - E_q|) I_q \\ W_p^b &= \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|E_p - E_q|) \end{aligned} \quad (12)$$

直观地认为，高斯滤波器的权重基于图像中位置  $p$  和  $q$  之间的空间距离和图像中  $p$  和  $q$  位置像素值的差。由于该滤波器的非线性特性，在实时应用受到限制。有些环节可以通过使用不同的方法来近似结果，以交互式速率进行计算。Durand 和 Dorsey 通过线性内插几个离散化的强度值来计算滤波器的响应，这些强度值在频域中用高斯内核滤波。Paris 和 Durand 将这项工作解释为一种高维线性卷积，其后是三线性插值和分裂，从而扩展了这项工作。这种想法的推广是使用双边网格，双边网格  $\Gamma$  是一个有规则取样的三维阵列，其中前两个维度定义图像的空间域，而第三个维度表示图像的范围。使用双边网格进行滤波操作，分为三个步骤进行：网格创建、处理和切片。为了处理双边网格，可以将三维函数应用其中，我们得到一个新的网格  $\Gamma' = f(\Gamma)$ 。对于双边滤波器， $f$  是高斯核的卷积，分别使用  $\sigma_s, \sigma_r$  作为域和距离维度上的方差。切片操作用于从双边网格中提取分段平滑的二维图像。

在双边网格中，使用图像  $E$  来执行交叉双边滤波以确定网格位置，同时存储来自数据图像  $I$  的像素值。切片操作随后使用边缘图像来恢复结果。Adams 等人通过统一单形有效地实现了 5D 彩色双边滤波，与 Paris 和 Durand 的方法相比，他们使用了更简单的插值操作。最近 Gastal 和 Oliveira 提出了一种计算适应信号的流形，这些流形之间的加权插值可以得到更高精度和更快速率的输出，并使用比之前方法更少的采样点数量。

Kovaleski 和 Oliveira 提出的方法通过使用双边网格生成一个亮度增强函数，网格使用两个不同的图片生成，其意义上更像是交叉双边滤波。但是切片操作通过使用第三幅图片来实现，其方法可以被写为：

$$bf(I) = s_B(G_{\sigma_s, \sigma_r} \otimes g_A(C)) \quad (13)$$

其中  $A, B, C$  都是单通道图片， $A$  中的每个像素都包含对应 RGB 通道的图像中的最大强度值，图像  $B$  包含图像  $I$  的亮度信息，图像  $C$  包含图像  $A$  中所有像素值超过某阈值  $t$  的所有像素点。进一步，RGB 通道中的最大强度值（图像  $A$ ）被用作检测饱和和像素的默认方式，正如我们将要展示的那样，这个选择会影响该算法处理过度曝光内容的能力。在生成图像  $A, B, C$  之前，图像使用 Gamma 2.2 曲线进行线性化处理， $g_A(C)$  是一个双边网格，使用图像  $I$  的边缘信息进行过滤，同时存储图像  $C$  的值。因此双边滤波公式可以被写为：

$$I_p^b = \frac{1}{W_p^b} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|B_p - A_q|) C_q$$

$$W_p^b = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|B_p - A_q|) \quad (14)$$

一个关键点在于，上式中使用最大强度值图像  $A$  倾向于导致 Kovaleski 和 Oliveira 的算法无法增强过度曝光的 SDR 图像的细节。通过使用最大强度信道值，他们的算法尽管在感知上存在差异，但他们对 RGB 三通道的信息平等处理，这种不一致反映在所得到的亮度增强函数中。这个关键点同样可以解决这个问题：用亮度图像  $B$  代替最大强度通道图像  $A$ ，这种操作不仅解决了曝光过度的图像的问题，而且极大地简化了亮度增强函数生成过程，将以上公式变为交叉双边滤波形式：

$$I_p^b = \frac{1}{W_p^b} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|B_p - B_q|) C_q$$

$$W_p^b = \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|B_p - B_q|) \quad (15)$$

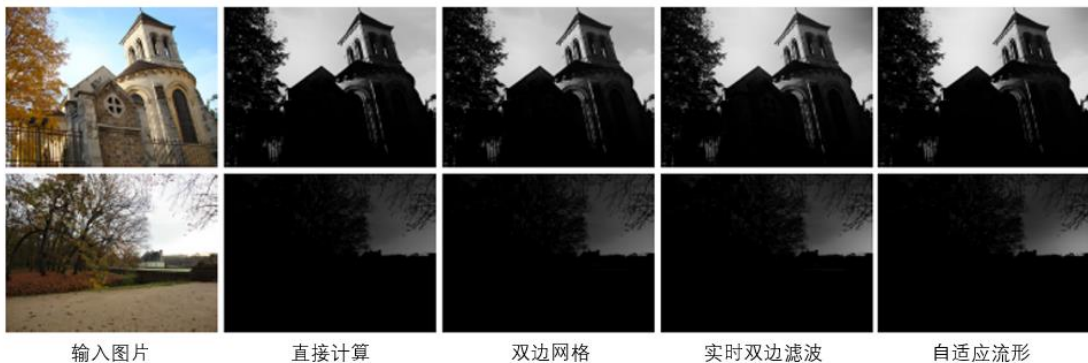


图 7-9 不同双向滤波算法的结果对比

公式(15)相比与之前公式来说，可以使用多种不同的双边滤波加速方法，如图 9 所示是使用不同种加速方法的效果，对比了不使用加速的双边滤波、双边网格、实时双边滤波和自适应流形四种结果。此外，滤波参数中本算法取  $\sigma_s = 150$ ，这时模糊滤波器的滤波结果不容

易被人类视觉系统察觉；同时，另一个滤波器参数  $\sigma_r = 25$  是通过大量实验测量得出，实验表明，这个取值更有利于得到亮度增强函数。

根据计算存储的亮度增强函数，作者使用和 Rempel 算法类似的步骤：输入图像  $I$  使用 Gamma 2.2 曲线进行线性化，使用上面公式计算得到亮度增强函数，随后将函数缩放至范围  $[1, \alpha]$  内，在实验中  $\alpha$  取值为 4。最后，线性化输入图像的值根据目标显示能力和选择的  $\alpha$  值进行性缩放，最低亮度为  $0.3\text{cd/m}^2$ ，最高亮度为  $1200\text{cd/m}^2$ 。然后通过重新调整的亮度增强函数将该缩放的图像逐点相乘来获得最终的 HDR 图像。

生成效果图如图 10 所示，结果图使用 DRIM 评价方法进行展现。



图 7-10 DRIM 对结果图进行评价

## 2. 基于视网膜反应的逆色调映射

Huo [24]在 2014 年提出一种基于人类视觉系统视网膜反应的逆色调映射算法，通过生理学上的研究，该方法可以避免出现很多现有方法都会发生的伪像问题。受到人类视觉系统特性的启发，这种逆色调映射方案的计算复杂度低，参数数量较少，并能获得高质量的 HDR 图像输出。

与真实世界场景的动态范围相比，人类视网膜的感光细胞具有较窄的动态范围，这要归功于人类视觉系统的适应机制，它使得眼睛首先适应一些亮度值，然后以相当小的动态感知图像范围围绕该亮度值。因此，人类视觉的基本过程是对整个场景的全局色调映射，这种全局功能可以用视网膜的反应和光强刺激之间的关系来描述：

$$\frac{R}{R_{\max}} = \frac{I^n}{(I^n + \sigma^n)} \quad (16)$$

其中  $R$  表示视网膜对光强  $I$  的刺激反应， $R_{\max}$  是反映的最大值， $\sigma$  是全局适应等级，它表示产生的相应强度是  $R_{\max}$  振幅的一半所需的强度，参数  $n$  是灵敏度控制指数。

实验证明，场景的绝对亮度信息对于人类视觉系统来说是次要的，通过亮度很定性的机制在视觉处理的非常早期阶段往往会被大量舍弃，而是用局部对比来传达广域场景的大量信息。因此，视网膜反应的更恰当描述应该是局部映射，局部自适应属性可以通过将全局自适



应水平  $\sigma$  改编为局部像素  $p$  的自适应水平  $\sigma_p$  来实现。设  $I_p$  表示像素  $p$  在实际场景中的辐射强度值，视网膜反应  $R_p$  被认为是眼睛感知到的 SDR 图像的强度值，可以表示为：

$$R_p = \frac{R_{\max} I_p^n}{(I_p^n + \sigma_p^n)} \quad (17)$$

对于逆色调映射， $I_p$  可以通过以下方式计算， $\delta$  是为了避免分母为 0：

$$I_p = \left( \frac{R_p \sigma_p^n}{(R_{\max} - R_p + \delta)} \right)^{1/n} \quad (18)$$

之前的研究认为灵敏度参数  $n$  通常取值在 0.7~1.0 之间，作者通过大量实验得出，取值为 0.9 时结果更好。 $\sigma_p$  是局部适应水平，它描述了一个像素邻域强度信息，通常可以使用算术平均值、几何平均值或图像局部区域的高斯模糊来确定该参数。这里，HDR 图像中像素  $p$  周围的局部强度  $I_{ps}$  用来作为  $\sigma_p$  的取值。在仅有 SDR 图像的情况下，基于 SDR 图像的最大亮度 255 被映射到 HDR 显示器的最大亮度的假设，算法首先计算 SDR 图像的  $L_{ps}$ ，随后将其乘以 HDR 显示器最大亮度与 255 的比值。

很多方法可以计算得到  $L_{ps}$ ，论文使用 Durand 和 Dorsey 提出的双边滤波，逆色调映射公式可以写为：

$$I_p = \left( \frac{R_p (I_{ps}(\sigma_m, \sigma_d))^n}{(R_{\max} - R_p + \delta)} \right)^{1/n} \quad (19)$$

输出的像素  $p$  的双边滤波  $L_{ps}$ ：

$$L_{ps}(\sigma_m, \sigma_d) = \frac{1}{W_p \sum_{q \in \Omega} f_{\sigma_m}(q-p) g_{\sigma_d}(R_q - R_p) R_q} \quad (20)$$

其中  $W_p$  是归一化因子：

$$W_p = \sum_{q \in \Omega} f_{\sigma_m}(q-p) g_{\sigma_d}(R_q - R_p) \quad (21)$$

其中  $\sigma_m$  表示高斯函数  $f$  在空间域的标准差：

$$f_{\sigma_m}(p | p = (x, y)) = K_m \exp \left\{ -\frac{(x^2 + y^2)}{\sigma_m^2} \right\} \quad (22)$$

其中  $\sigma_d$  表示高斯函数  $g$  在亮度域下的标准差， $K_m$  是归一化因子， $\Omega$  表示整幅图像。 $\sigma_m$  和  $\sigma_d$  分别取值为 16 和 0.3。如图 10 所示为方法计算结果与其他两种方法的对比。

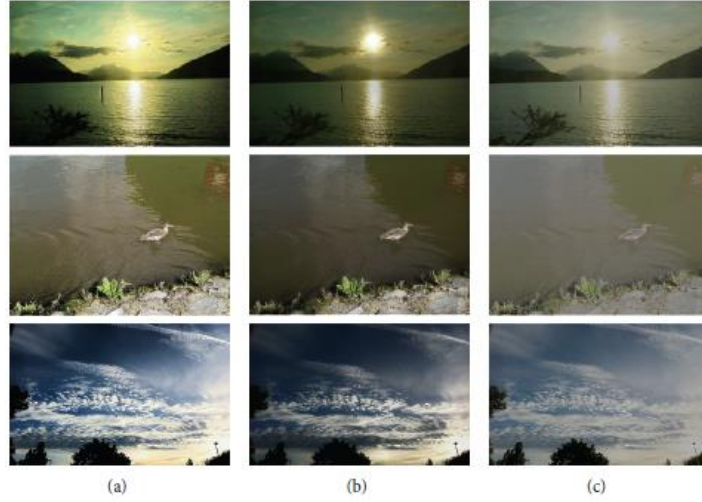


图 7-11 算法效果图和与其他方法的对比：(a) 论文方法，(b) iPG，(c) SDR2HDR

### 3. 基于深度学习的逆色调映射

Endo [23]针对逆色调映射中过度曝光和曝光过少、细节损失的情形进行优化，提出了一种基于深度学习的方法，其网络基于卷积神经网络，将输入图像转化为多幅不同曝光度的图片，再组合为一张 HDR 图像输出。

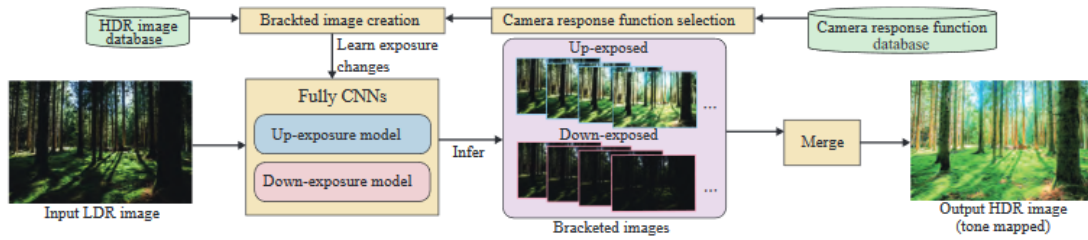


Fig. 2. Overview of the proposed method. The flow is decomposed into the learning and inference phases. In the learning phase, the bracketed LDR images are first created from HDR databases by simulating cameras. Next, we let our fully CNNs learn the changes in the exposures of the bracketed images. In the inference phase, the learned CNNs compute LDR images with different exposures from a single input LDR image. The final HDR image is then generated from these bracketed LDR images.

图 7-12 Endo 提出的方法流程图

12 所示是方法的流程框图。首先，作者通过使用不同的非线性相机反应曲线来改变图像的曝光度，由此形成了多幅不同曝光度的图片，数据集就是以 SDR 图像为输入，多幅不同曝光度的图片为输出组成。其计算方法如下：

$$Z_{i,j} = f(E_i \Delta t_j)$$

$$\Delta t_j = \frac{1}{\tau^{T/2}}, \dots, \frac{1}{\tau^2}, \frac{1}{\tau}, 1, \tau, \tau^2, \dots, \tau^{T/2}$$
(23)

其中  $Z_{i,j}$  表示 SDR 图像对应点  $i$  在曝光指数  $j$  下的像素值， $f$  表示相机反应函数， $E_i$  表示 HDR 图像在点  $i$  处的像素值。 $\Delta t_j$  为曝光时间，作者通过指数定义，实验中采用  $T=8, \tau=\sqrt{2}$ 。随后，在使用得到的多张不同曝光度的照片，通过合并算法进行计算，从而得到一张完整的 HDR 图片。



Endo 使用的是编解码器结构，并在对应层间有跳跃式连接。先将一张 SDR 图片编码至深度网络，这里通过 2D 卷积神经网络实现，随后再将图像的深度语义特征解码为不同曝光度的 SDR 图片，这时使用的是 3D 卷积神经网络。这种网络用于将图像的曝光度提高或降低，从而得到多幅图片用于之后合成 HDR 图片。

在具体网络结构方面，编码器网络有 9 层卷积核大小为 4 的卷积层，并包含批归一化和 Leaky-ReLU。解码器网络使用 3D 去卷积神经网络，其中前三层的卷积核为  $4 \times 4 \times 4$ ，其余层为  $3 \times 4 \times 4$ 。类似地，出最后一层外，其他层都包含批归一化和 Leaky-ReLU，而最后一层的激活函数为 Sigmoid。

在这个编解码网络中，解码器使用完全编码的向量，这意味着整个图像的潜在特征都包含在内，为了将输入图像中的局部和底层信息逐步融入解码器，作者在编解码器的对应层之间添加了跳跃式连接，具体网络结构图如图所示：

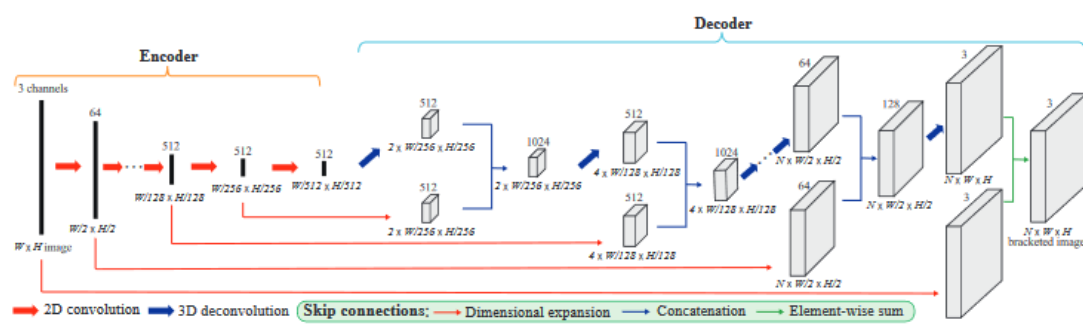


Fig. 5. Our network architecture. The same architecture is used both for up- and down-exposure models.

图 7-13 网络结构图

通过长达一个月的训练，取得了满意效果。其与一些传统算法的比如下图所示。

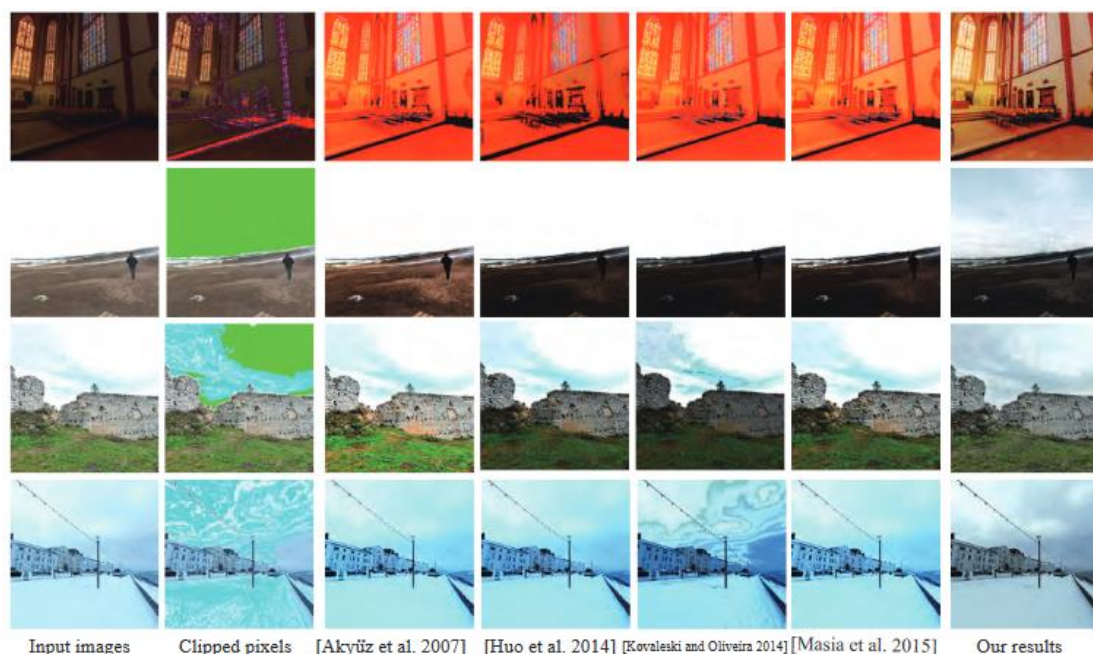


Fig. 6. Tone-mapped results of HDR images generated by each rTMO. To visualize the clipped pixels in the input LDR images, pixels of the images in the second column are painted in red/green if they are completely under/overexposed (i.e., black/white) or in magenta/cyan if one or two channels are under/overexposed, respectively. While there is amplified noise in the dark regions of other results, our results exhibit natural tone reproduction with the least noise. Please enlarge images in the electronic version.

图 7-14 Endo 算法效果图和与其他方法的对比

#### 4. 基于生成对抗网络的逆色调映射模型

S. Ning 等人[3]认为由于逆色调映射过程相当于是向低动态范围的图像中添加暗区和亮区的细节，希望可以达到一个信息增加的过程，因此采用生成对抗网络的对抗机制可以很好的完成生成任务。生成对抗网络是由 I. Goodfellow 提出的基于博弈的模型，其模型本身由两个网络构成，一个是生成网络，在图像任务中用于生成图像，实现图像的转换任务；另一个是判别网络，用于判别输入给判别网络的图像是由生成器生成的还是源图像。生成对抗网络的训练过程就是，通过训练生成网络使得生成的图像越来越接近源图像，从而可以“蒙骗”判别网络进行错误的决策；训练判别网络的过程就是使判别网络可以更准确的分辨出生成的图像和源图像，从而使得这两部分网络的训练过程成为一种博弈式对抗过程，期望的训练结果是能够得到一个稳态。

算法上讲，GAN 的判别网络可以看成是将图像数据映射到判别概率的函数  $D: D(x) \rightarrow (0,1)$ ，对于一个固定的生成器  $G$ ，判别器  $D$  可能被训练用于粪便图像是来自训练数据（概率接近 1）还是来自生成器（概率接近 0）。这种方法的优点是其分类边界灵活，使用纯概率的方法可以得到挖掘信息更深的模型，这种模型可以清晰的判别出多类别或某一类与其他类之间存在的差异特征。

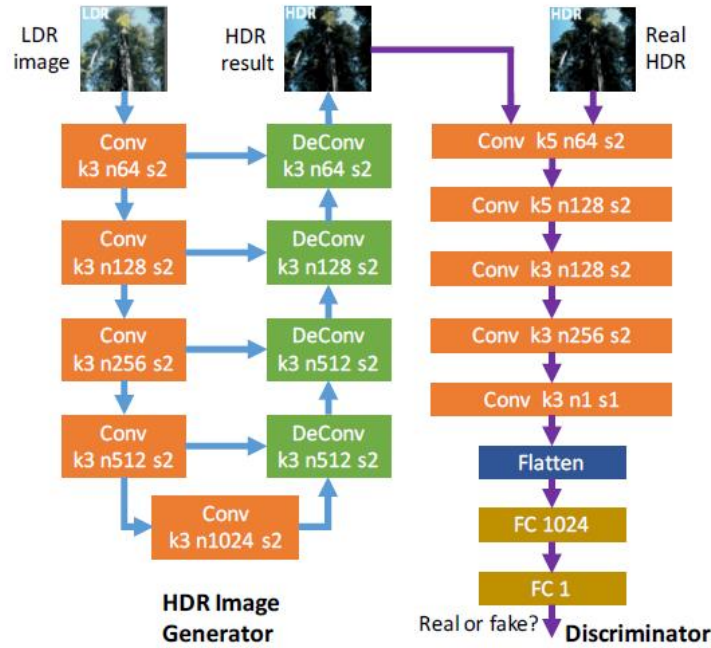


图 7-15 网络结构图

作者在论文提出了一种 iTMN 可以用于将 SDR 转换为 HDR 的生成对抗网络，网络结构如图 15 所示，生成网络采用 U-Net 类似结构，这是因为 U-Net 可以在图像重构的过程中同时使用深层信息和浅层信息一步一步地重建图像，因此 U-Net 会很适合本任务。用于约束网络的损失函数是一种混合损失函数，包括生成对抗网络定义相关的对抗损失函数和生成网络内容相关的内容损失函数：

$$\min_G \max_D \underbrace{\lambda L_G(L, H)}_{\text{content loss}} + \underbrace{R_{G,D}(L, H)}_{\text{adversarial regularizer}} \quad (24)$$

其中内容损失函数包含两个部分，分别是均方差和差分均方差，希望通过均方差从像素值方面上约束网络，通过差分均方差在周围像素点的关系上进行约束：

$$E_{(L,H) \sim p_{data}} \left[ \|G(L) - H\|_F^2 + \alpha \left( \|d_x G(L) - d_x H\|_F^2 + \|d_y G(L) - d_y H\|_F^2 \right) \right] \quad (25)$$

除此之外，对抗损失函数即由 GAN 网络的定义得出，可以被归纳为：

$$E_{H \sim p_{data}} [\log(1 - D(H))] + E_{L \sim p_{data}} [\log(D(G(L)))] \quad (26)$$

数据使用 HDR 视频切片单帧图片，经过多个色调映射算法处理后选取效果最好的 SDR 图像与原生 HDR 图像组成数据集，训练使用的训练集包含 2660 对图像，测试集包含 140 对图像，实现采用 TensorFlow 框架。训练后网络在测试数据集上做验证，评测方法采用 HDR-VDP-2、mPSNR 和 SSIM，结果与 Huo[24]、KO[22] 和 DrTM0[23] 做对比，并同时与提出结构的两种变式做对比，其中 NoDMSE 是提出的结构中除去 L2 部分，NoAdvReg 是提出结构除去对抗部分，即使用单独的 U-Net，评价结果如下：

**Table 1: Comparison for various methods**

| Method    | iTMN   | NoDMSE | NoAdvReg | Huo [2] | KO [3] | DrTM [4] |
|-----------|--------|--------|----------|---------|--------|----------|
| HDR-VDP-2 | 73.77  | 72.47  | 70.52    | 60.84   | 56.45  | 73.21    |
| mPSNR     | 21.61  | 20.62  | 20.07    | 20.64   | 14.25  | 18.73    |
| SSIM      | 0.8098 | 0.7298 | 0.7242   | 0.7707  | 0.6459 | 0.7968   |

该方法在三种不同的评价指标效果均有提升，说明该算法在对比度提升、图像内容增强和图像结构信息上都有一定的进步。除此之外，也对网络的参数进行了一些实验，关于是否需要运用 GAN 模型在发色调映射任务上的问题，从之前的数据和图 16 所示的效果图中都显示，使用 GAN 可以得到更好的转换效果。

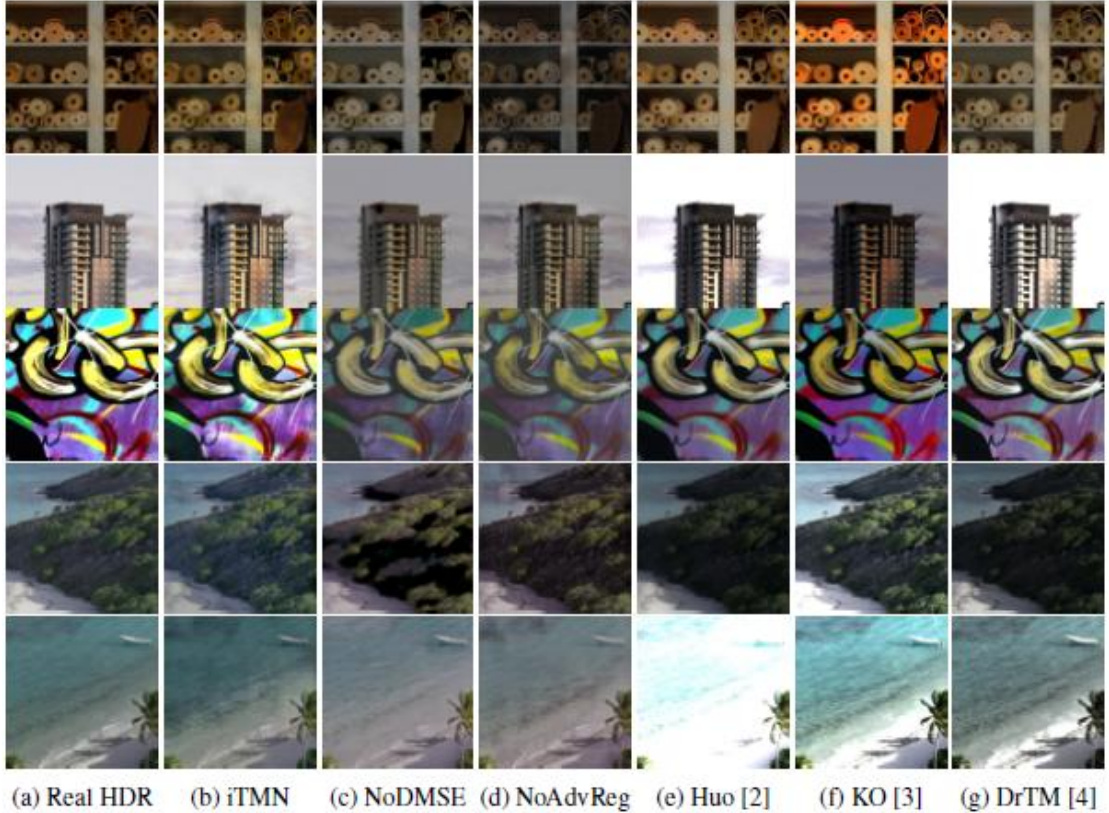


图 7-16 S. Ning 算法效果图和与其他方法的对比



## 5. 利用深度卷积网络的单次曝光 HDR 重建

相机的感光元件只能一时间捕捉范围有限的亮度，因此传统方法中很多都是使用多个曝光度的照片来合成 HDR 图片。Eilertsen 在文章[27]中提出了一种使用单次曝光的图像来进行 HDR 图像重建，本质是在亮度饱和区域上预测丢失的信息，这种任务很适合用深度学习的算法来解决，因此算法是基于一种深度卷积网络（CNN）来预测丢失的 HDR 信息。

算法包括以下步骤：

首先，算法需要对输入的 SDR 图片进行线性化、颜色校正等预处理。线性化方法中，为了降低复杂度而使用了一种平均的相机响应函数来线性化图像；颜色校正上，也是使用固定的白平衡转换来简化预处理的过程；量化恢复若使用 CNN 来完成则更接近超分辨率等工作，因此本方法中并未进行量化恢复的处理。

随后，算法的模型使用自动编解码器结构实现，对应层也使用了跳跃式连接，但与 U-Net 不同的是，网络的解码器部分是在图像的对数域进行处理，因此跳跃式连接还需要将图像的特征转换到对数域再与解码器的各层输出结合，可以通过公式(27)表示：

$$\bar{h}_i^D = \sigma \left( W \left[ \log \left( f^{-1} \left( h_i^E \right) \right) + \varepsilon \right] + b \right) \quad (27)$$

其中  $h_i^E$  和  $h_i^D$  分别表示编码网络和解码网络对应层的输出， $f^{-1}$  是将图片线性化的处理函数，随后的操作就是将两部分信息组合起来作为混合信息送给下一卷积层处理。网络的结构如图 17 所示：

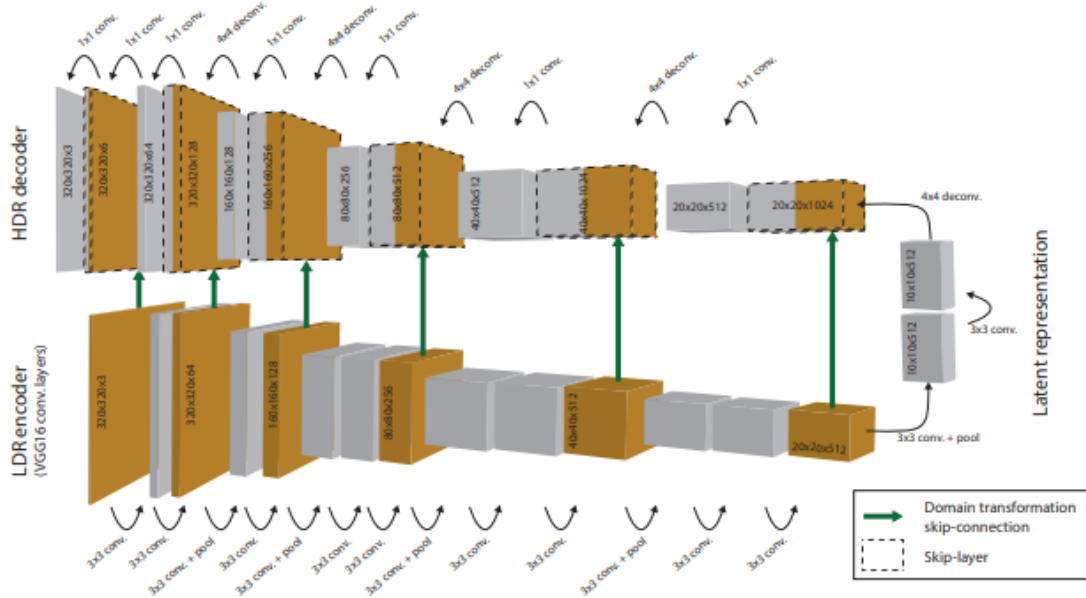


图 7-17 Eilertsen 提出方法的网络结构

最终输出的 HDR 图像也是由网络输出的对数域图片和输入图片线性化后加权得到，可以由公式(28)表示：

$$H_{i,c} = (1 - \alpha_i) f^{-1}(D_{i,c}) + \alpha_i \exp(\hat{y}_{i,c}) \quad (28)$$

其中  $D_{i,c}$  表示输入的 SDR 图像， $\hat{y}_{i,c}$  表示网络输出的对数域的 HDR 图像， $i, c$  分别表示图像

的标号和颜色通道， $\alpha_i$  是用户自定义取值，用于调整两种图像来源的权重。

最后，网络的损失函数是通过将图像分解为亮度分量（Illumination）和漫反射分量（Reflectance），用于得到图像不同维度的信息，然后对两种分量转换至对数域后，在 L2 范数下进行约束，最终的损失函数由两个分量在对数域上的 L2 范数差的加权和来确定，可以由公式 (29) 表示：

$$L(\hat{y}, H) = \frac{\lambda}{N} \sum_i \left| \alpha_i \left( \log(I_i^{\hat{y}}) - \log(I_i^y) \right) \right|^2 + \frac{1-\lambda}{3N} \sum_{i,c} \left| \alpha_i \left( \log(R_{i,c}^{\hat{y}}) - \log(R_{i,c}^y) \right) \right|^2 \quad (29)$$

其中  $I, R$  分别表示图像的亮度分量和漫反射分量， $\hat{y}, y$  分别表示网络的输出图片和真实图片， $\lambda$  是用户指定的权值参数。训练使用的是预训练的模型，实验表明经过预训练的模型可以重建出局部亮度更高的 HDR 结果，这与真实 HDR 图像更为接近，随后对网络使用常用参数在与 Endo[23] 相同方法下构建的数据集进行训练，训练结果如图 18 所示，可以看出得到的模型能够较好的重建 HDR 图像的高亮部分，这与真实的 HDR 图片很接近。



图 7-18 Eilertsen 使用的深度学习方法的重建结果

#### 参考文献：

- [1] H. Landis. Production-Ready Global Illumination, in: SIGGRAPH Course Notes 16, pp. 87 - 101, 2002.
- [2] A.O. Akyuz, R. Fleming, B.E. Riecke, E. Reinhard. Do HDR displays support LDR content?: a psychophysical evaluation, in: ACM Transactions on Graphics, Vol.26, no 3 p. 38, 2007.
- [3] S. Ning, H. Xu, L. Song, R. Xie, W. Zhang. Learning an Inverse Tone Mapping Network with a Generative Adversarial Regularizer, in: arXiv preprint arXiv:1804.07677 (2018).
- [4] F. Durand, J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images, in: ACM transactions on graphics (TOG). Vol. 21. No. 3, 2002.
- [5] G.W. Larson, H. Rushmeier, C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes, in: IEEE Transactions on Visualization & Computer Graphics, No. 4, pp. 291-306, 1997.

- [6] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda. Photographic tone reproduction for digital images, in: ACM transactions on graphics. Vol. 21. No. 3: 267–276, 2002.
- [7] B. Masia, S. Agustin, R.W. Fleming, O. Sorkine, D. Gutierrez. Evaluation of reverse tone mapping through varying exposure conditions, in: ACM transactions on graphics (TOG). Vol. 28. No. 5, pp. 160, 2009.
- [8] A.G. Rempel, M. Trentacoste, H. Seetzen, H.D. Young, W. Heidrich. Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs, in: ACM transactions on graphics (TOG). Vol. 26. No. 3, pp. 39, 2007.
- [9] L. Meylan, D. Scott, S. Sabine. The reproduction of specular highlights on high dynamic range displays, in: Color and Imaging Conference. Vol. 2006. No. 1, pp. 333–338, 2006.
- [10] L. Meylan, D. Scott, S. Sabine. Tone mapping for high dynamic range displays, in: Human Vision and Electronic Imaging XII. Vol. 6492, pp. 649210, 2007.
- [11] P. Didyk, R. Mantiuk, M. Hein, H.P. Seidel. Enhancement of bright video features for HDR displays, in: Computer Graphics Forum. Vol. 27, No. 4, pp. 1265–1274, 2008.
- [12] V. Vapnik. The nature of statistical learning theory. Springer science & business media, 2013.
- [13] F. Banterle, P. Ledda, K. Debattista, A. Chalmers. Inverse tone mapping, in: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia. ACM, pp. 349–356, 2006.
- [14] F. Banterle, P. Ledda, K. Debattista, A. Chalmers. A framework for inverse tone mapping, in: The Visual Computer Vol. 23, No. 7, pp. 467–478, 2007.
- [15] F. Banterle, P. Ledda, K. Debattista, A. Chalmers. A GPU-friendly method for high dynamic range texture compression using inverse tone mapping, in: Proceedings of graphics interface, pp. 41–48, 2008.
- [16] F. Banterle, P. Ledda, K. Debattista, A. Chalmers. Expanding low dynamic range videos for high dynamic range applications, in: Proceedings of the 24th Spring Conference on Computer Graphics. ACM, pp. 33–41, 2008.
- [17] S. Lin, L. Zhang. Determining the radiometric response function from a single grayscale image, in: 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–26 June 2005.
- [18] S. Lin, J. Gu, S. Yamazaki, H.Y. Shum. Radiometric Calibration from a Single Image, in: Proceedings of the 2004 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 938–945. Washington, DC, USA, 27 June – 2 July 2004.
- [19] R. Mantiuk, S. Daly, K. Myszkowski, H. Seidel. Predicting Visible Differences in High Dynamic Range Images—Model and Its Calibration, in: Human Vision and Electronic Imaging X, IST SPIE’ s 17th Annual Symposium on Electronic Imaging, pp. 204–214, 2005
- [20] R. Kovalski, M. Oliveira. High-Quality Brightness Enhancement Functions for Real-Time Reverse Tone Mapping, in: The Visual Computer. Vol. 25 No. 5–7, pp. 539–547, 2009.
- [21] J. Chen, S. Paris, F. Durand. Real-Time Edge-Aware Image Processing with

- the Bilateral Grid, in: ACM Transactions on Graph. Vol. 26, No. 3, pp. 103, 2007.
- [22] R. Kowaleski, M. Oliveira. High-quality reverse tone mapping for a wide range of exposures, in: Graphics, Patterns and Images (SIBGRAPI), Vancouver, Canada, 12-14 August 2014.
- [23] Y. Endo, Y Kanamori, J Mitani. Deep reverse tone mapping, in: ACM Transaction on Graph Vol. 36 No. 6, 2017.
- [24] Y. Huo, F. Yang, L. Dong, V. Brost. Physiological inverse tone mapping based on retina response, in: The Visual Computer, Vol. 30 No. 5, pp. 507-517, 2014.
- [25] ITU-R. BT.2087: Colour conversion from Recommendation ITU-R BT.709 to Recommendation ITU-R BT.2020, in: International Telecommunication Union-Radiocommunication Sector, 2015. Available: <http://www.itu.int/rec/R-REC-BT.2087-0-201510-I>
- [26] ITU-R. BT.2407: Colour gamut conversion from Recommendation ITU-R BT. 2020 to Recommendation ITU-R BT. 709, in: International Telecommunication Union-Radiocommunication Sector, 2017. Available: <https://www.itu.int/pub/R-REP-BT.2407-2017>
- [27] Gabriel Eilertsen, Joel Kronander, Gyorgy Denes, Rafał K. Mantiuk, Jonas Unger. HDR image reconstruction from a single exposure using deep CNNs. In: ACM Transactions on Graphics (Proc. of SIGGRAPH Asia 2017), 36(6), Article 178, 2017.